Regularity lemmas in a Banach space setting

Guus Regts[*]

June 24, 2015

Abstract

Szemerédi's regularity lemma is a fundamental tool in extremal graph theory, theoretical computer science and combinatorial number theory. Lovász and Szegedy [19] gave a Hilbert space interpretation of the lemma and an interpretation in terms of compactness of the space of graph limits. In this paper we prove several compactness results in a Banach space setting, generalising results of Lovász and Szegedy [19] as well as a result of Borgs, Chayes, Cohn and Zhao [5].

1 Introduction

1.1 The regularity lemma

Szemerédi's regularity lemma [27] is a fundamental tool in extremal graph theory, theoretical computer science and combinatorial number theory. See [18] for a survey. The lemma has many interpretations, variations and extensions. See for example [17, 12, 1, 25, 11, 13, 9, 28, 19, 26, 5].

Very roughly the lemma says something of the form: for each $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that the vertex set of any graph can be partitioned into at most k parts, such that for 'almost' all pairs of parts the edges between that pair of parts behaves 'almost' like a random bipartite graph, where 'almost' depends on ε. The weak regularity lemma of Frieze and Kannan [12] weakens the requirements of the partition in the regularity lemma and measures the error of approximation with respect to the cut norm. This has as a consequence that the constant k can be taken to be much smaller. From the perspective of the adjacency matrix of a graph this means that one approximates this matrix with a bounded sum of cut matrices (in particular this gives a low rank approximation) such that their difference is small with respect to the cut norm. This is exactly the point of view we take in this paper: we want to find various types of low rank approximations to matrices and tensors, when measured in a particular norm.

Our work is inspired by the work of Lovász and Szegedy [19] and Borgs, Chayes, Cohn and Zhao [5] relating the compactness of the space of graph limits to Szemerédi's regularity lemma. We refer to the book by Lovász [21] for more details on graph limits. In [20] Lovász and Szegedy used the weak version of the regularity lemma [12] to assign a limit object to a convergent sequence of dense graphs. This limit object is no longer a graph, but a symmetric measurable function $W : [0,1]^2 \to [0,1]$, called a *graphon*. In [19] Lovász and

---

[*]University of Amsterdam. Email: `guusregts@gmail.com`.

Szegedy showed that the space of graphons, equipped with the cut metric is compact, interpreting this result as an analytical form of the regularity lemma. Their compactness result implies various kinds of regularity lemmas varying from weak to very strong. It has recently been extended by Borgs, Chayes, Cohn and Zhao [5] to the space of $\mathbb{R}$-valued functions W with bounded p-norm, the L^p-*graphon space* (for any fixed $p > 1$).

1.2 Compactness

We will now describe the compactness of the graphon space, which is denoted by $\mathcal{W}$, more precisely (for details concerning definitions we refer to the next section), after which we mention some of the results in the present paper.

Let $W : [0,1]^2 \to \mathbb{R}$ be an integrable function. Consider for $p, q \in [1, \infty]$, W as a kernel operator $W : L^p([0,1]) \to L^q([0,1])$; that is, for $f \in L^p([0,1])$, $(Wf)(x) := \int_{[0,1]} W(x,y)f(y)d\lambda$. The $p \mapsto q$-operator norm is defined by

$$\|W\|_{p \mapsto q} = \sup_{\|f\|_p = 1} \|Wf\|_q,$$

where $\| \cdot \|_s$ denotes the s-norm on the space $L^s([0,1])$. The norm $\| \cdot \|_{\infty \mapsto 1}$ is equivalent to the cut norm. Define an equivalence relation $\sim$ on $\mathcal{W}$ as follows: $W \sim W'$ if for each $\varepsilon > 0$ there exists a measure preserving bijection $\tau : [0,1] \to [0,1]$ such that $\|W - \tau W'\|_{\infty \mapsto 1} \leq \varepsilon$ for $W, W' \in \mathcal{W}$. Then the result of Lovász and Szegedy [19] can be stated as follows:

$$\text{the space } (\mathcal{W}, \| \cdot \|_{\infty \mapsto 1}) / \sim \text{ is compact.} \tag{1}$$

The result of Borgs, Chayes, Cohn and Zhao [5] then says that we can replace $\mathcal{W}$ with the symmetric functions in $L^p([0,1]^2)$ of norm at most 1 for any fixed $p > 1$.

In this paper we will show that in (1) we can also replace the norm $\| \cdot \|_{\infty \mapsto 1}$ by the norm $\| \cdot \|_{q \mapsto \frac{q}{q-1}}$, if we replace $\mathcal{W}$ by the unit ball of $L^p([0,1]^2)$, provided that $p > \frac{q}{q-1}$, cf. Theorem 4.1. In fact, we generalise (1), replacing the space $\mathcal{W}$ by a special weakly compact subset of a Banach space X, the relation $\sim$ by an equivalence relation obtained from a subgroup of the group of autmorphisms of X and the norm $\| \cdot \|_{\infty \mapsto 1}$ by an operator-type norm, cf. Theorem 3.2. From this result it is then easy to derive the results of Borgs, Chayes, Cohn and Zhao. In Section 4 we will also utilise it to include $q \mapsto \frac{q}{q-1}$-norms and apply it to higher order tensors. In Section 5 we will apply it to ℓ^p-spaces.

Our method is based on work of the author and Schrijver [24]. In [24] the compactness result of Lovász and Szegedy was extended to a general Hilbert space setting, putting emphasis on the possibility of using different norms than the cut norm and the use of groups and moreover using a different method of proof. Consequently, our proof of Theorem 3.2 does not use the martingale convergence theorem. Thus it yields a different proof of the compactness result of Borgs, Chayes, Cohn and Zhao [5]. However, there are some similarities. To prove Theorem 3.2 we need a result from [24], cf. Lemma 3.5, which may be viewed as a weak regularity lemma in a Hilbert space setting, and which generalises weak regularity results from [12, 11].

1.3 Algorithms and applications

Some of the existing versions of the weak and strong regularity lemmas come with efficient algorithms for finding a low rank approximation (or regularity partition). These algorithms

have been applied to find approximation schemes for various sorts of dense instances of counting and optimisation problems [12, 11] and to property and parameter testing [1, 6, 7]; see also the book by Lovász [21].

Some of our results can also be put to algorithmic use. In particular, in the ℓ^p setting, Lemma 5.3 can be used to give approximation algorithms for certain instances of computing the matrix $p \mapsto q$ norm and for finding approximate Nash-equilibria in two player games, in a similar spirit as has been done by Barman in [3]. In the L^p setting sampling algorithms from Borgs, Chayes, Lovász, Sós and Vesztergombi [7] can be applied to find low rank approximations to matrices yielding polynomial time approximation algorithms for computing the matrix $p \mapsto q$ norm for dense matrices. This is work in progress and we will report on it in a forthcoming paper [23].

1.4 Organisation

In the next section we will discuss some preliminaries and set up some notation. In Section 3 we will state and prove Theorem 3.2, the aforementioned generalisation of (1). We will also deduce some consequences from it. In Section 4 we will apply this theorem to L^p spaces and in Section 5 to ℓ^p spaces.

2 Preliminaries and notation

In this section we will give some preliminaries on Lebesgue spaces and set up some notation. We refer to [8] for functional analytic background and to [15] for measure theoretic background.

Lebesgue spaces For a measure space $(\Omega, \mathcal{A}, \mu)$ and $p \in [1, \infty]$ we denote by $L^p(\Omega)$ the linear space of equivalence classes of μ-integrable complex (or real) valued functions $f : \Omega \to \mathbb{C}$ (or $\mathbb{R}$) with bounded p-norm, which is defined as

$$\|f\|_p := \left(\int |f|^p d\mu \right)^{1/p} \text{ for } p < \infty, \text{ and}$$

$$\|f\|_\infty := \inf_{t \geq 0}\{|f(x)| \leq t \mid \text{ for } \mu\text{-almost all } x\}.$$

(Recall that two functions are equivalent if they are equal μ-almost everywhere.)

For our results it often does not matter whether we use real or complex-valued functions. So we generally do not distinguish between the complex and real-valued cases. If however we want to specify that we work over the field of real numbers, we denote this by $L^p_{\mathbb{R}}(\Omega)$. We often omit the reference to the sigma-algebra $\mathcal{A}$ and the measure μ. In case $\Omega = [0,1]^l$ for some $l \in \mathbb{N}$ we will always equip it with the Borel (or Lebesgue) sigma algebra and with the Lebsegue measure λ.

For a set Ω, and $p \in [1, \infty]$, $\ell^p(\Omega)$ is just equal to $L^p(\Omega)$ with $\mathcal{A}$ the power set of Ω and μ the counting measure. We often write L^p and ℓ^p whenever the underlying space is clear.

For a normed space $(Y, \|\cdot\|)$ we denote its closed unit ball by $B(Y, \|\cdot\|)$, which is defined as $\{y \in Y \mid \|y\| \leq 1\}$. Often we just write $B(Y)$. Let (Ω, μ) be a probability space, i.e., $\mu(\Omega) = 1$. An important property of the space $L^p(\Omega)$, that we will often use, is the nesting of the closed unit balls: for any $1 < p < q < \infty$ we have

$$B(L^\infty) \subset B(L^q) \subset B(L^p) \subset B(L^1). \tag{2}$$

For the closed unit balls in $l^p(\Omega)$, for any set Ω, the opposite inclusions hold:

$$B(\ell^1) \subset B(\ell^p) \subset B(\ell^q) \subset B(\ell^\infty). \tag{3}$$

Weak topologies We often need weak topologies in this paper. So it wil be convenient to introduce some notation. Let $(X, \|\cdot\|)$ be a normed space. Let X^* be the dual space of X, i.e., the space of all continuous linear functions $f : X \to \mathbb{C}$ (or $\mathbb{R}$). Let $Y \subset X^*$ be a subset of X^*. The *weak topology induced by Y* on X is generated by sets of the form $B_{x_0,f,\varepsilon} = \{x \in X \mid |f(x) - f(x_0)| < \varepsilon\}$ for $x_0 \in X$, $f \in Y$ and $\varepsilon > 0$. If we do not specify Y we mean the weak topology induced by X^*.

For $p \in [1, \infty]$ let

$$p^* := \begin{cases} \frac{p}{p-1} & \text{if } 1 < p < \infty, \\ 1 & \text{if } p = \infty, \\ \infty & \text{if } p = 1. \end{cases}$$

It is well known that L^{p^*} isometrically embeds into $(L^p)^*$ and this embedding is onto if $p < \infty$ (which holds for $p = 1$ provided that μ is sigma finite). So if Ω is a probability space, we can consider the weak topology on $B(L^p)$ induced by L^q for each $q \geq p^*$ (by (2)). We denote this by $(B(L^p), w_q)$. For ℓ^p we need to take $q \leq p^*$. This is then denoted by $(B(\ell^p), w_q)$.

Group actions If a group G acts on a set S this induces an action on the functions from S to $\mathbb{C}$ (or $\mathbb{R}$) via $gf(s) := f(g^{-1}s)$ for $g \in G$, a function f and $s \in S$. Moreover, G has a natural action on S^l for any $l \in \mathbb{N}$. In particular, if $(\Omega, \mathcal{A}, \mu)$ is a measure space and a group G acts on Ω such that $\mu(gA) = \mu(A)$ for all measurable sets $A \in \mathcal{A}$ and $g \in G$ (in which case we call elements of G *measure preserving bijections*), then G acts on $X = L^p(\Omega^l)$ for any p and l and preserves the p-norm. If a group acts on a set S and $T \subset S$ we denote by GT the set $\{gt \mid g \in G, t \in T\}$. We call $T \subseteq S$ *G-stable* if $GT = T$.

3 Compact orbit spaces in Banach spaces

In this section we will state and prove our main results concerning compact orbit spaces in Banach spaces and discuss some consequences, which may be viewed as regularity lemmas in a Banach space setting.

3.1 The main theorems

Before we can state our results, we need some definitions. Let $X = (X, \|\cdot\|)$ be a normed space and let R be a bounded subset of X^*. We define a seminorm $\|\cdot\|_R$ and pseudo metric d_R on X by

$$\|x\|_R := \sup_{r \in R} |r(x)| \qquad d_R(x,y) := \|x - y\|_R \tag{4}$$

for $x, y \in X$.

For a pseudo metric space (X, d) let $\mathrm{Aut}(X)$ denote the group of invertible maps $g : X \to X$ that preserve d. Let G be a subgroup of $\mathrm{Aut}(X)$. Define a pseudo metric d/G on X by

$$(d/G)(x,y) := \inf_{g \in G} d(x, gy)$$

for $x, y \in X$. Note that since $(d/G)(x, y)$ is just equal to the distance between the G-orbits of x and y, this implies that d/G is indeed a pseudo metric. For our purposes it is sometimes more convenient to work with $(X, d/G)$ than with X/G, but note that $(X, d/G)$ is compact if and only if X/G is compact. Recall that a (pseudo) metric space is called *totally bounded* if for each $\varepsilon > 0$ it can be covered with finitely many balls of radius ε.

We can now state our first result about compactness of orbit spaces in Banach spaces, which we will prove in Section 3.2.

Theorem 3.1. *Let $(X, \|\cdot\|)$ be a Banach space and let G be a subgroup of $\mathrm{Aut}(X)$. Let $R \subseteq B(X^*)$ be G-stable, and let $W \subset X$ be G-stable and weakly compact. If $(W, d_R/G)$ is totally bounded, then $(W, d_R/G)$ is compact.*

Showing that $(W, d_R/G)$ is totally bounded may not be very simple. However, if W is somehow 'close' to a Hilbert space (as will be made precise below) and R can be embedded into a bounded subset of this Hilbert space, then totally boundedness of $(W, d_R/G)$ can be deduced from compactness of the space of sums of elements from R modulo G.

For a subset Y of a linear space X and $k \in \mathbb{N}$ we define

$$k \cdot Y := \{y_1 + \ldots + y_k \mid y_i \in Y\}.$$

Note that when Y is convex, $k \cdot Y$ is just equal to kY. Let X be a normed space and let H be a Hilbert space. We call $W \subset X$ *H-small*[1], if here exists a contractive linear map $T : H \to X$ and a function $c : (0, \infty) \to \mathbb{N}$ such that $W \subset c(\varepsilon)T(B(H)) + \varepsilon B(X)$ for each $\varepsilon > 0$. Note that T gives rise to a contractive linear map $T^* : X^* \to H^*$, *the adjoint of T*, defined by $f \mapsto (h \mapsto f(T(h)))$ for $f \in X^*$ and $h \in H$. Identifying H with H^* (which we will always do) this gives a contractive linear map $TT^* : X^* \to X$. When we talk about a H-small space we will implicitly assume the presence of the maps T, T^* and c.

Theorem 3.2. *Let $(X, \|\cdot\|)$ be a Banach space and let G be a subgroup of $\mathrm{Aut}(X)$. Let $R \subseteq B(X^*)$ be G-stable and let $W \subset X$ be G-stable and weakly compact. If there exists a Hilbert space H such that W is H-small and if $(k \cdot (TT^*(R)), d_R/G)$ is compact for each $k \in \mathbb{N}$, then $(W, d_R/G)$ is compact.*

Observe that when X is a Hilbert space, Theorem 3.2 reduces to [24, Theorem 2.1].

Before we give a proof of Theorem 3.2, let us remark that the compactness of the L^p-graphon space proved by Borgs, Chayes, Cohn and Zhao [5] follows almost immediately from it. Let $R \subset L^\infty([0,1]^2)$ be defined by

$$R := \{\chi_{A \times B} \mid A, B \subseteq [0,1] \text{ measurable}\}.$$

This makes $\|\cdot\|_R$ into the cut norm. Let us denote the group of measure preserving bijections $\tau : [0,1] \to [0,1]$ by $S_{[0,1]}$. Then $d_R/S_{[0,1]}$ is equal to $\delta_\square$, the cut metric. The result of Borgs, Chayes, Cohn and Zhao is then equivalent to the following:

Corollary 3.3. *Let $p > 1$. Then $(B(L^p([0,1]^2)), d_R/S_{[0,1]})$ is compact.*

---

[1]The notion of *H-smallness* can be seen as a qualitative refinement of relative weak compactness, by a classical result of Grothendieck [14]. Moreover, related notions have been used elsewhere in functional analysis, cf. [10].

Proof. In case $p \geq 2$ we take $X = L^2$, otherwise we set $X = L^1$. Write $W = B(L^p) \subset X$ and let $H = L^2$. Then H is a Hilbert space and W is H-small by [5], or see Lemma 4.4 (the map T being the identity.) Note also that T^*T restricted to R is the identity. Since any measurable set A can be mapped onto any interval of length $\lambda(A)$ by a measure preserving bijection, cf. [22], it follows that $(R^k, d_R/S_{[0,1]})$ is compact and hence $(k \cdot R, d_R/S_{[0,1]})$ as well, as it is a continuous image of $(R^k, d_R/S_{[0,1]})$ (see [24]). Let us first assume that $p < \infty$. Then, as $(B(L^p), w_{p*})$ is compact by the Banach-Alaoglu theorem, it follows that $(B(L^p), w_\infty)$ is compact. (Since w_∞ induces a weaker topology than w_{p*}.) In case $p = \infty$, we may use that $B(L^\infty) \subset B(L^s)$, for any fixed $s > 1$, is a closed convex set with respect to the s-norm. This implies that $(B(L^\infty), w_{s*})$ is compact and hence $(B(L^\infty), w_\infty)$ is compact, as desired. Theorem 3.2 now implies that $(B(L^p), d_R/G)$ is compact. □

With little additional effort we can derive something similar as Corollary C.8 in [5]:

Corollary 3.4. *Given a function* $\kappa : (0, \infty) \to (0, \infty)$, *let*

$$W_\kappa := \bigcap_{\varepsilon > 0} \kappa(\varepsilon) B(L_\mathbb{R}^\infty([0,1]^2)) + \varepsilon B(L_\mathbb{R}^1([0,1]^2)).$$

Then $(W_\kappa, d_R/S_{[0,1]})$ *is compact.*

Proof. Throughout the proof we assume that we work over the real numbers. Let again $H = L^2$ and $X = L^1$. Then $W_\kappa \subset X$ is clearly H-small by (2). So by Theorem 3.2 it suffices to check that (W_κ, w_∞) is compact. We may assume that W_κ is contained in $B(L^1)$. Since $(B(L^\infty)^*, w_\infty)$ is compact by the Banach-Alaoglu theorem and contains $B(L^1)$, it suffices to show that W_κ is weakly closed.

To this end let μ be an element of the weak closure of W_κ. Then μ is a finitely additive signed measure on $[0,1]^2$. We will show that μ is in fact sigma additive. To see this, let $\varepsilon > 0$ and choose $\delta = \varepsilon\kappa(\varepsilon)^{-1}$. Then for each measurable set $U \subset [0,1]^2$ there exists $w \in W_\kappa$ such that $|\mu(U) - \int_U w d\lambda| \leq \varepsilon$. This implies that if $\lambda(U) < \delta$, then $\mu(U) \leq 3\varepsilon$, as for each $w \in W_\kappa$, $|\int_U w d\lambda| \leq \lambda(U)\kappa(\varepsilon) + \varepsilon = 2\varepsilon$. Let now $(U_n)_{n \in \mathbb{N}}$ be a collection of pairwise disjoint measurable sets. Write $U = \cup_{n \in \mathbb{N}} U_n$. Then, by sigma additivity of the Lebesgue measure, we have $\lambda(\cup_{n > N} U_n) \to 0$, as N tends to infinity. This implies that $\mu(\cup_{n > N} U_n) \to 0$, as N tends to infinity. So $\mu(U) = \sum_{n \in \mathbb{N}} \mu(U_n)$, showing that μ is sigma additive.

Since μ is finite, the Radon-Nikodym theorem implies the existence of an integrable function f such that $\int_U f d\lambda = \mu(U)$ for each measurable U.

We will now show that f is again contained in W_κ. Let $\varepsilon > 0$ and define $U := \{x \mid f(x) > \kappa(\varepsilon)\}$, $V := \{x \mid f(x) < -\kappa(\varepsilon)\}$ and let $\delta > 0$. Then there exists $w \in W_\kappa$ such that $|\int_U (f - w) d\lambda| \leq \delta$ and $|\int_V (f - w) d\lambda| \leq \delta$. This implies

$$\int_{U \cup V} |f| d\lambda = \int_U f d\lambda - \int_V f d\lambda \leq \int_U w d\lambda - \int_V w d\lambda + 2\delta$$

$$\leq \int_{U \cup V} |w| d\lambda + 2\delta \leq \lambda(U \cup V)\kappa(\varepsilon) + \varepsilon + 2\delta. \tag{5}$$

Since (5) holds for any $\delta > 0$, this shows that $f \in W\kappa$ and finishes the proof. □

Remark 1. The requirement that we work over the real numbers in the corollary above can be omitted if we choose to define W_κ in the following way:

$$W_\kappa := \bigcap_{\varepsilon > 0} (\kappa)(\varepsilon) \big(B(L_\mathbb{R}^\infty([0,1]^2)) + iB(L_\mathbb{R}^\infty([0,1]^2))\big) + \varepsilon\big(B(L_\mathbb{R}^1([0,1]^2)) + iB(L_\mathbb{R}^1([0,1]^2))\big).$$

3.2 Proofs of Theorem 3.1 and 3.2

The proofs of Theorems 3.1 and 3.2 follow the same pattern as the proof of Theorem 2.1 in [24]. Let us start with the proof of Theorem 3.1.

Proof of Theorem 3.1. Since $(W, d_R/G)$ is totally bounded it suffices to show that $(W, d_R/G)$ is complete. Let $a_1, a_2, \ldots$ be a Cauchy sequence in $(W, d_R/G)$. We may assume that $(d_R/G)(a_i, a_{i+1}) < 2^{-i}$ for each i. Next find iteratively $g \in G$ such that $d_R(a_i, ga_{i+1}) < 2^{-i}$ and replace a_{i+1} by ga_{i+1}. Then $d_R(a_i, a_j) < 2^{-i+1}$ for all i and $j \geq i$ and hence this sequence is Cauchy with respect to d_R. We will show that this sequence is also convergent.

Let for $n \in \mathbb{N}$, A_n denote the weak closure of the set $\{a_n, a_{n+1}, \ldots\}$. By weak compactness, $A := \cap_{n \in \mathbb{N}} A_n$ is not empty. Let $a \in A$. We will show that a_n converges to a with respect to d_R. Let $\varepsilon > 0$. Choose N such that for all $n, m > N$, $d_R(a_n, a_m) \leq \varepsilon$. We will show that $d_R(a_n, a) \leq 2\varepsilon$ for all $n > N$. To this end fix $r \in R$. Then there exists $m \geq N$ such that $|(r, a_m - a)| \leq \varepsilon$. Then for any $n > N$ we have

$$|(r, a - a_n)| \leq |r, a - a_m)| + |(r, a_m - a_n)| \leq 2\varepsilon,$$

showing that $d_R(a_n, a) \leq 2\varepsilon$. Hence (a_n) converges to a with respect to d_R. □

To prove Theorem 3.2, we need the following lemma from [24], which may be viewed as a weak regularity lemma in a Hilbert space setting. In particular, choosing appropriate R, we recover results from [12] and [11] respectively.

Lemma 3.5 ([24]). *Let H be any Hilbert space. Let $R \subseteq B(H)$ be closed under multiplication by elements of $\{z \in \mathbb{C} \mid |z| \leq 1\}$ (by elements of $[-1, 1]$ if H is real). Then for any $k \in \mathbb{N}$,*

$$B(H) \subseteq k \cdot R + \frac{1}{\sqrt{k}} B(H, \| \cdot \|_R).$$

For convenience of the reader we will give a proof.

Proof. Let us denote the inner product on H by $\langle \cdot, \cdot \rangle$. Let $a \in B(H)$. Write $a_0 = a$. If $\|a_i\|_R > \frac{1}{\sqrt{k}}$ for some $i \geq 0$, then there exists $r \in R$ such that $|\langle a_i, r \rangle| > \frac{1}{\sqrt{k}}$. Set $a_{i+1} = a_i - \langle a_i, r \rangle r$. Note that by induction we have that $a_{i+1} - a_0 \in i \cdot R \subset k \cdot R$, as $|\langle a_i, r \rangle| \leq 1$ by Cauchy-Schwarz. Then

$$\|a_{i+1}\|_2^2 = \|a_i\|_2^2 - 2|\langle a_i, r \rangle|^2 + |\langle a_i, r \rangle|^2 \|r\|_2^2 = \|a_i\|_2^2 - |\langle a_i, r \rangle|^2 (2 - \|r\|_2^2) < \|a_i\|_2^2 - 1/k,$$

by definition of r and the fact that $R \subseteq B(H)$. By induction we have that $\|a_{i+1}\|_2^2 < 1 - i/k$. This shows that for some $i \leq k$ we must have $\|a_i\|_R \leq \frac{1}{\sqrt{k}}$ and hence finishes the proof. □

We can now prove Theorem 3.2.

Proof of Theorem 3.2. By Theorem 3.1 it suffices to show that $(W, d_R/G)$ is totally bounded. We may assume that R is closed under multiplication by elements of $\{z \in \mathbb{C} \mid |z| \leq 1\}$ (elements of $[-1, 1]$ if our spaces are real), as this does not changes the norm induced by R. Neither does it change the compactness of $(k \cdot (TT^*(R)), d_R/G)$ for any k. To show that $(W, d_R/G)$ is totally bounded, choose $\varepsilon > 0$. By assumption there exists a Hilbert space H such that W is H-small. This means there exists a contractive linear map $T : H \to X$ and a constant $c > 0$ such that $W \subset cT(B(H)) + \varepsilon B(X)$. As $(k \cdot (TT^*(R)), d_R/G)$ is compact there

7

exists a finite set $F \subset k \cdot (TT^*(R))$ such that $k \cdot (TT^*(R)) \subset GF + \frac{\varepsilon}{c}B(X, \|\cdot\|_R)$. Setting $k := \lceil c^2\varepsilon^{-2} \rceil$, we have by Lemma 3.5, since $T(B(H, \|\cdot\|_{T^*(R)})) \subseteq B(X, \|\cdot\|_R)$,

$$W \subseteq cT(B(H)) + \varepsilon B(X) \subset c(k \cdot (TT^*(R)) + k^{-1/2}T(B(H), \|\cdot\|_{T^*(R)})) + \varepsilon B(X) \subseteq$$
$$c(GF + \frac{\varepsilon}{c}B(X, \|\cdot\|_R) + k^{-1/2}B(X, \|\cdot\|_R)) + \varepsilon B(X) \subseteq GcF + 3\varepsilon B(X, \|\cdot\|_R).$$

This shows that $(W, d_R/G)$ is totally bounded and finishes the proof. $\qquad\square$

Remark 2. The proof shows that if R is already closed under multiplication by elements of $\{z \in \mathbb{C} \mid |z| \le 1\}$, then it is enough to demand that $(k \cdot TT^*(R), d_R/G)$ is merely totally bounded.

3.3 Weak and strong regularity

The following is implicit in the proof of Theorem 3.2, and may be viewed as an extension of the weak regularity lemma from the Hilbert space setting to a Banach space setting.

Lemma 3.6. *Let $(X, \|\cdot\|)$ be a Banach space, let $W \subset X$ and let $R \subseteq B(X^*)$. Suppose there exist a Hilbert space H such that W is H-small. Then for each $\varepsilon > 0$ there exists $k \le \lceil c(\varepsilon)/\varepsilon \rceil^2$ such that for each $w \in W$ there exists $\alpha \in \{z \in \mathbb{C} \mid |z| \le 1\}^l$ and $x_1, \ldots, x_l \in TT^*(R)$, with $l \le k$, such that $\|w - c \sum_{i=1}^l \alpha_i x_i\|_R \le 2\varepsilon$.*

In [19], Lovász and Szegedy applied the compactness of the graphon space, cf. (1), to derive an approximation result for graphons, cf. [19, Lemma 5.2]. This result implies several types of regularity lemmas varying from the weak regularity lemma [12], to the original lemma [27], to a 'super strong' variant [1]. See [21] for more details. We can derive something similar in our Banach space setting:

Lemma 3.7. *Let $(X, \|\cdot\|)$ be a Banach space, let $G \subseteq \mathrm{Aut}(X)$, let $R \subseteq B(X^*)$ and let $W \subset X$ be G-stable and suppose that $(W, d_R/G)$ is compact. Let for $k \in \mathbb{N}$, $Y_k \subset W$ be G-stable such that $Y := \cup_{k \in \mathbb{N}} Y_k$ is dense in W (w.r.t. $\|\cdot\|$). Let $h : (0, \infty) \times \mathbb{N} \to (0, 1)$ be any function. Then for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for any $w \in W$ there exists $w' \in W$ and $y \in Y_m$, with $m \le N$, such that*

$$\|w - w'\|_R \le h(\varepsilon, m) \quad and \quad \|w' - y\| \le \varepsilon.$$

The proof goes along the same lines as the proof of Lemma 5.2 in [19].

Proof. We may assume that h is monotonically decreasing in its second variable. Let $\varepsilon > 0$. Since Y is dense in W, for each $w \in W$, there exists $n \in \mathbb{N}$ and $y \in Y_n$ such that $\|w - y\| \le \varepsilon$. Let $f(w)$ denote the smallest n such that there exists $y \in Y_n$ with $\|w - y\| \le \varepsilon$. For $w \in W$ let $O(w)$ denote the d_R/G-open set defined as

$$O(w) := \{w' \in W \mid (d_R/G)(w, w') < h(\varepsilon, f(w))\}.$$

By compactness of $(W, d_R/G)$, it follows that there exists $w_1, \ldots, w_t$ such that the union of the $O(w_i)$ contains W. This means that for any $w \in W$ there exists $i \in \{1, \ldots, t\}$ and $y \in Y_m$, with $m = f(w_i)$, such that $\|w_i - y\| \le \varepsilon$ and $(d_R/G)(w, w_i) < h(\varepsilon, f(w_i))$. Choosing $N := \max_{i=1,\ldots,t} f(w_i)$ and a suitable $g \in G$ applied to both w_i and y we arrive at the desired conclusion. $\qquad\square$

8

Since the norm $\|\cdot\|$ on X satisfies $\|x\| \leq \|x\|_R$ for each $x \in X$, taking $h(\varepsilon, m) = \varepsilon$ for all m, Lemma 3.7 implies several types of weak regularity lemmas by taking different choices of Y_k. In particular, it allows to use different types of approximations not based on sums of elements from $\{z \in \mathbb{C} \mid |z| \leq 1\} TT^*(R)$ as in Lemma 3.6. We will elaborate a bit more on this in the next two sections, where we will apply the results obtained here to the spaces $\ell^p(\mathbb{N}^l)$ and $L^p([0,1]^l)$.

4 Compact orbit spaces in $L^s(([0,1]^l)$

In this section we will apply Theorems 3.1 and 3.2 to the Banach space $X = L^s([0,1]^l)$ for some fixed $s \in [1, \infty)$ and $l \in \mathbb{N}$. Let for any $q \in [1, \infty]$,

$$R_l^q := \{r_1 \otimes \ldots \otimes r_l \mid r_i \in B(L^q([0,1]))\} \subset B(L^q([0,1]^l)).$$

Recall that $S_{[0,1]}$ denotes the group of measure preserving bijections $\tau : [0,1] \to [0,1]$. We will derive the following result from Theorems 3.1 and 3.2, generalising Corollary 3.3, the compactness result of Borgs, Chayes, Cohn and Zhao [5].

Theorem 4.1. *Let $p \in (1, \infty]$ and let q be such that $q > p^*$. Fix $l \in \mathbb{N}$. Then the space $(B(L^p([0,1]^l)), d_{R_l^q})/S_{[0,1]}$ is compact.*

We will prove Theorem 4.1 in Section 4.1 below. Let us first make some remarks and state some consequences.

Borg, Chayes, Cohn and Zhao [5], who proved Theorem 4.1 for $l = 2$ and $q = \infty$, already noted that the same result cannot be true when $p = 1$. We note here that in case $p = 2$ we really need that $q > 2$ since $B(L^2([0,1]^2), d_{R_2^2})/S_{[0,1]}$ is not totally bounded. To see this take f_i to be the function which is constant i on the rectangle $[0, 1/i] \times [0, 1/i]$. Since for rank 2 functions f we have $\|f\|_{R_2^2} \leq \|f\|_2 \leq 2\|f\|_{R_2^2}$, we may as well replace $d_{R_2^2}$ by d_2. Then for any $\pi \in S_{[0,1]}$ we have for $i = 2^k < j = 2^l$ with $l \geq k + 2$,

$$\int_{[0,1]^2} |f_i - \pi f_j|^2 \geq \int_{[0,1]^2} |f_i - f_j|^2 \geq \frac{(2^k - 2^l)^2}{2^l} \geq 1 - 2^{k+1-l} \geq 1/2,$$

implying that we need to take $q > 2$ when $p = 2$. This example of course generalises to any $l > 2$ and p, q such that $q = p^*$.

Janson [16] showed, using the Riesz-Thorin theorem, that in case $p = \infty$ (and $l = 2$), Theorem 4.1 remains true if one replaces the $\|\cdot\|_{R_2^q}$ norm by the $p_0 \mapsto q_0$ operator norm for any $p_0, q_0 > 1$. See [16, Lemma E.6]. We can also use the Riesz-Thorin theorem in our setting for $l = 2$. (We need to work over the complex numbers though. This does not cause any problems for real-valued functions, as the complex and real operator norms are within a constant factor of each other.) Let us take $p > 1$ and $q > p^*$. Taking $f \in B(L^p([0,1]^2))$, we then have $\|f\|_{p^* \mapsto p} \leq 1$. Define for $\theta \in (0,1)$ p_θ and q_θ by

$$\frac{1}{p_\theta} := \frac{1-\theta}{p^*} + \frac{\theta}{q} \quad \text{and} \quad \frac{1}{q_\theta} := \frac{1-\theta}{p} + \frac{\theta}{q^*}. \tag{6}$$

Then by the Riesz-Thorin theorem (cf. [4, Theorem 1.1.1]) $\|f\|_{p_\theta \mapsto q_\theta} \leq \|f\|_{q \mapsto q^*}^\theta = \|f\|_{R_l^q}^\theta$. By Theorem 4.1 this implies the following:

Corollary 4.2. *Let $p, q \in (1, \infty]$ such that $q > p^*$. Let $\theta \in (0, 1)$ and let p_θ and q_θ be defined as in (6). Then $(B(L^p([0,1]^2)), \| \cdot \|_{p_\theta \mapsto q_\theta})/S_{[0,1]}$ is compact.*

By Lemma 3.7 the following is a direct corollary to Theorem 4.1 and the previous corollary.

Corollary 4.3. *Let $p, q \in (1, \infty]$ such that $q > p^*$. Fix $l \in \mathbb{N}$. Let for $k \in \mathbb{N}$, $F_k \subset B(L^p([0,1]^l))$ be $S_{[0,1]}$-stable such that $F := \cup_{k \in \mathbb{N}} F_k$ is a dense subset of $B(L^p([0,1]^l))$ (w.r.t. $(\| \cdot \|_{q^*})$. Let $h : (0, \infty) \times \mathbb{N} \to (0, 1)$ be any function. Then for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for any $g \in B(L^p([0,1]^l))$ there exists $g' \in B(L^p([0,1]^l))$ and $f \in F_m$, with $m \leq N$, such that*

$$\|g - g'\|_{R_l^q} \leq h(\varepsilon, m) \quad and \quad \|g' - f\|_{q^*} \leq \varepsilon. \tag{7}$$

In case $l = 2$ and p_θ and q_θ are defined by (6), the R_l^q norm may be replaced by $\| \cdot \|_{p_\theta \mapsto q_\theta}$ in (7).

There are various choices for the sets F_k in the corollary above. For example one can take F_k to be those function that are sums of at most k rectangles (a *rectangle* is a function of the form $c\chi_{A_1 \times \cdots \times A_l}$ with A_i measurable and c a constant.) A special case is to take F_k to be those sums of rectangles that are coming from a partition of $[0, 1]$ into at most k sets. Another choice would be to take F_k to be those functions that take only k different values (i.e. step functions with k steps).

4.1 Proof of Theorem 4.1

To prove Theorem 4.1, we need some preliminary results.

The next lemma implies that $B(L^p([0,1]^l)) \subset L^1([0,1]^l)$ is H-small for $H = L^2([0,1]^l)$ and any $p \in (1, 2]$.

Lemma 4.4. *Let (Ω, μ) be any probability space. Let $p > p' \geq 1$ and $\varepsilon > 0$. Then there exists a constant C such that $B(L^p(\Omega)) \subseteq CB(L^\infty(\Omega)) + \varepsilon B(L^{p'}(\Omega))$.*

Proof. Define for convenience $B^s := B(L^s(\Omega))$, for any $s \in [1, \infty]$. Fix any $K \geq 1$ and let $f \in B^p$. Let $A := \{\omega \in \Omega \mid |f(\omega)| > K\}$. Then $f = f_1 + f_2$ with $f_1 = f\chi_{\Omega \setminus A}$ and $f_2 = f\chi_A$. Clearly, $f_1 \in KB^\infty$. Next we consider $\|f_2\|_{p'}$:

$$\|f_2\|_{p'}^{p'} = \int_\Omega |f_2|^{p'} d\mu \leq \int_\Omega |f_2|^{p'} (|f|/K)^{p-p'} d\mu \leq K^{p'-p} \int_\Omega |f|^p d\mu \leq K^{p'-p}.$$

Hence choosing C in such a way that $C^{1-p/p'} = \varepsilon$ we obtain the desired result. $\qquad \square$

Lemma 4.5. *Let $p > s \geq 1$ and let $\varepsilon > 0$. Then for any $k \in \mathbb{N}$ there exists a constant C such that $R_k^p \subseteq CR_k^\infty + \varepsilon R_k^s$.*

Proof. The proof is by induction on k. The case $k = 1$ follows by Lemma 4.4. Now suppose $k > 1$. Choose any $p' \in (s, p)$ and let $\delta > 0$. By induction, there exists a constant C_1 such that $R_{k-1}^p \subseteq C_1 R_{k-1}^\infty + \delta R_{k-1}^{p'}$ and $R_1^p \subseteq C_1 R_1^\infty + \delta R_1^{p'}$. Let us define for $s, t \in [1, \infty]$,

$$R_{k-1,1}^{s,t} := \{f_1 \otimes f_2 \mid f_1 \in R_{k-1}^s, f_2 \in R_1^t\}.$$

Then

$$R_k^p \subseteq C_1^2(R_k^\infty + C_1 \delta(R_{k-1,1}^{\infty,p'} + R_{k-1,1}^{p',\infty}) + \delta^2 R_k^{p'}. \tag{8}$$

Next choose $\eta > 0$. Then by induction there is a constant C_2 such that $R_1^{p'} \subseteq C_2 R_1^\infty + \eta R_1^s$ and $R_{k-1}^{p'} \subseteq C_2 R_{k-1}^\infty + \eta R_{k-1}^s$. Plugging this into (8) we obtain that

$$R_k^p \subseteq CR_k^\infty + C_1\eta\delta(R_{k-1,1}^{\infty,s} + R_{k-1,1}^{s,\infty}) + \delta^2 R_k^{p'},$$

where $C := C_1^2 + 2\delta C_1 C_2$. Finally, using that $R_l^t \subseteq R_l^s$ for any $t \geq s$ and $l \in \mathbb{N}$, and first choosing δ such that $\delta^2 \leq 1/3\varepsilon$ and then η such that $C_1\delta\eta \leq 1/3\varepsilon$, we obtain the desired inclusion. $\qquad\square$

Proof of Theorem 4.1. Let us write R^s instead of R_l^s for any $s \in [1,\infty]$. Let us denote $S_{[0,1]}$ by G. Define $X = L^{q^*}([0,1]^l)$. Then $W = B(L^p([0,1]^l)) \subset X$, as $p > q^*$. The proof now proceeds in a number of steps.

Let

$$R := \{\chi_{A_1} \times \cdots \times \chi_{A_l} \mid A_i \subset [0,1] \text{ measurable for each } i\}.$$

Note that $\|\cdot\|_R$ and $\|\cdot\|_{R^\infty}$ define equivalent norms on $L^1([0,1]^l)$. More precisely, for $f \in L^1([0,1]^l)$, we have

$$\|f\|_R \leq \|f\|_{R^\infty} \leq 4^l\|f\|_R,$$

since for the norm $\|\cdot\|_{R^\infty}$ we may restrict to functions $r = r_1 \otimes \cdots \otimes r_l$ in (4) such that the r_i only take values $-i, i, -1$ and 1. This implies that $(B(L^\infty), d_{R^\infty}/G)$ is compact by Corollary 3.3, or more precisely by the proof of Corollary 3.3, as the proof did not depend on l being equal to 2.

We will next show:

$$(B(L^\infty), d_{R^q}/G) \text{ is compact.} \tag{9}$$

To do so, let (f_n) be a sequence in $B(L^\infty)$. By compactness of $(B(L^\infty), d_{R^\infty}/G)$ it has a convergent subsequence (which we may assume to be (f_n) itself) that converges to some $f \in B(L^\infty)$ with respect to d_{R^∞}/G. So it suffices to show that (f_n) converges to f with respect to d_{R^q}/G. Let $\varepsilon > 0$. Let C be the constant from Lemma 4.5 such that $R^q \subseteq \varepsilon R^1 + CR^\infty$. Fix $N \in \mathbb{N}$ such that $(d_{R^\infty}/G)(f_n - f) < \varepsilon/C$ for all $n \geq N$. In other words, for each $n \geq N$ there exists $g_n \in G$ such that $\|g_n f_n - f\|_{R^\infty} < \varepsilon/C$. Then, writing $f_n' = g_n f_n$,

$$\|f_n' - f\|_{R^q} = \sup_{r \in R^q} r(f_n' - f) \leq \sup_{\substack{r_1 \in CR^\infty \\ r_2 \in \varepsilon R^1}} r_1(f_n' - f) + r_2(f_n' - f)$$

$$\leq C\|f_n' - f\|_{R^\infty} + \varepsilon\|f_n' - f\|_{R^1} \leq \varepsilon + 2\varepsilon = 3\varepsilon,$$

showing that f_n converges to f with respect to d_{R^q}/G. This proves (9). We will now show:

$$(B(L^p), d_{R^q}/G) \text{ is totally bounded in } X. \tag{10}$$

To this end let $\varepsilon > 0$ and let K be the constant from Lemma 4.4 such that $B(L^p) \subseteq KB(L^\infty) + \varepsilon B(L^{q^*})$ and note that $B(L^{q^*}) = B(X) \subseteq B(X, \|\cdot\|_{R^q})$. Write $\varepsilon' = \varepsilon/K$. By compactness of $(B(L^\infty), d_{R^q}/G)$, there exists a finite set $F \subset B(L^\infty)$ such that $B(L^\infty) \subseteq GF + \varepsilon' B(L^\infty, \|\cdot\|_{R^q})$. Then

$$B(L^p) \subseteq KGF + K\varepsilon' B(L^\infty, \|\cdot\|_{R^q}) + \varepsilon B(X, \|\cdot\|_{R^q}) \subseteq KGF + 2\varepsilon B(X, \|\cdot\|_{R^q}),$$

proving (10).

By the Banach-Alaoglu theorem we know that $(B(L^p), w_{p^*})$ is compact (here we assume that $p < \infty$, as this case is already covered). This implies that $(B(L^p), w_q)$ is compact, since w_q yields a weaker topology than w_{p^*}, as $q > p^*$. This finishes the proof by Theorem 3.1. $\quad\square$

5 Compact orbit spaces in $\ell^q(\mathbb{N}^l)$

In this section we will apply Theorem 3.1 to the Banach space $X = \ell^q(\mathbb{N}^l)$ for $q \in (1, \infty]$. Let $S_\mathbb{N}$ be the group of invertible maps $\tau : \mathbb{N} \to \mathbb{N}$. We denote the metric induced by $\| \cdot \|_p$ by d_p for any $p \in [1, \infty]$. We will derive the following result:

Theorem 5.1. *Let $p \in [1, \infty)$ and let $q > p$. Fix $l \in \mathbb{N}$. Then the space $B(\ell^p(\mathbb{N}^l)), d_q)/S_\mathbb{N}$ is compact.*

We will prove this result in Section 5.1 below. Let us first note that we really need $p < q$. Lemma 5.3 below is not true for $p = q$ and is a direct corollary of Theorem 5.1. It seems that one can only take $q = p$ when both of them are equal to 2, but then one has to use a different (weaker) metric and the orthogonal group instead of $S_\mathbb{N}$, cf. [24].

By Lemma 3.7 the following is a direct corollary to Theorem 5.1:

Corollary 5.2. *Let $p \in [1, \infty)$ and let $q > p$. Fix $l \in \mathbb{N}$. Let for $k \in \mathbb{N}$, $X_k \subset B(\ell^p(\mathbb{N}^l))$ be $S_\mathbb{N}$-stable such that $X := \cup_{k \in \mathbb{N}} X_k$ is a dense subset of $B(\ell^p(\mathbb{N}^l))$ (w.r.t. $\| \cdot \|_q$). Then for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for any $y \in B(\ell^p(\mathbb{N}^l))$ there exists $x \in X_m$, with $m \leq N$, such that $\|y - x\|_q \leq \varepsilon$.*

We can take $X_k := (k \cdot R) \cap B(\ell^p(\mathbb{N}^l))$ with $R = \{r_1 \otimes \cdots \otimes r_l \mid r_i \in B(\ell^p(\mathbb{N}))\}$ in the corollary above. Then the corollary says that for each $\varepsilon > 0$ there is N such that each $y \in B(\ell^q(\mathbb{N}^l))$ can be approximated in the q-norm by a tensor of rank at most N. Unfortunately, Corollary 5.2 does not give any explicit bounds on N. It is not to be expected that the bounds on N will be much better than the bound given by Lemma 5.3 below (at least for $q = \infty$). Indeed, for $q = \infty$, Alon, Lee, Schraibman and Vempala [2] showed that for any $n \times n$ Hadamard matrix M and $\varepsilon \in (0, 1)$ one has that for any matrix M', such that $\|M - M'\|_\infty \leq \varepsilon$, the rank of M' is at least $(1 - \varepsilon^2)n$. Since all entries of a Hadamard matrix are 1 or -1, we need to take $p = \Omega(\log n)$, to make sure that $\|M\|_p$ is bounded. The bound given by Lemma 5.3 on N is then of order $\varepsilon^{-\log n}$, which is polynomial in n.

5.1 Proof of Theorem 5.1

For any $q \in [1, \infty]$, a subset $K \subset S$ and $x \in \ell^q(S)$ we define $x_K \in \ell^q(S)$ by

$$x_K := \begin{cases} x(i) \text{ if } i \in K, \\ 0 \text{ otherwise.} \end{cases}$$

We need the following lemma for the proof of Theorem 5.1.

Lemma 5.3. *Let S be any set. Let $p \in [1, \infty)$ and let $q > p$. For any $\varepsilon > 0$ there exists $k = k(p, q, \varepsilon) \in \mathbb{N}$ such that for any $x \in B(\ell^p(S))$ there exists a set $K \subset S$ of size at most k such that $\|x_K - x\|_q \leq \varepsilon$. In case $q = \infty$ we can take $k = \varepsilon^{-p}$.*

Proof. Let $C \subseteq S$ be a countable set of points, which we identify with $\mathbb{N}$, such that $x(i) = 0$ for $i \notin C$. We may assume that x satisfies $|x(1)| \geq |x(2)| \geq \ldots$. Since $\|x\|_p \leq 1$, it follows that $|x(i)|^p \leq 1/i$ for all i. Let $K = \{1, \ldots, k\}$ for some $k \in \mathbb{N}$ to be fixed later. Then for $q < \infty$,

$$\|x_K - x\|_q^q = \sum_{i \notin K} |x(i)|^q \leq \sum_{i > k} (1/i)^{q/p}. \tag{11}$$

It is a well known fact that the series $\sum_{i=1}^{\infty}(1/i)^s$ converges for every $s > 1$. This implies that for k large enough, (11) will be bounded by ε^q, which finishes the first part of the proof.

In case $q = \infty$, we have

$$\|x_K - x\|_\infty = \sup_{i > k} |x(i)| \le (1/i)^{1/p}.$$

So taking k such that $(1/k)^{1/p} \le \varepsilon$, we are done. $\qquad\square$

Note that Lemma 5.3 is not true if one replaces q by p. One can simply take $x(i) = n^{-1}$ for $i = 1, \ldots, n^p$ and $x(i) = 0$ for $i > n^p$. Then for any constant k not depending on n and any set $K \subset \mathbb{N}$ of size k we have $\|x_K - x\|_p^p \le (n^p - k)n^{-p} = 1 - kn^{-p} = 1 - o(1)$.

Despite the fact that Lemma 5.3 is really simple, it can be utilised for algorithmic purposes, since we can go over all bounded size subsets of $\{1, \ldots, n\}$ in time polynomial in n. This is used in [23] to give approximation algorithms for computing matrix $p \mapsto q$ norms.

Now we can give a proof of Theorem 5.1.

Proof of Theorem 5.1. Write $W = B(\ell^p(\mathbb{N}^l))$. We may assume that $q < \infty$ since $d_\infty / S_\mathbb{N}$ yields a coarser topology than $d_{q'} / S_\mathbb{N}$ for any $q' < \infty$. Let now $R = B(\ell^{q^*}(\mathbb{N}^l))$. Then $d_R = d_q$.

If $p > 1$, then by the Banach-Alaoglu Theorem (W, w_{p^*}) is compact, which implies that (W, w_{q^*}) is compact since w_{q^*} induces a weaker topology than w_{p^*}, as $\infty > q > p$ and hence $1 < q^* < p^*$. If $p = 1$, we use the fact that $B(\ell^1(\mathbb{N}^l))$ is a closed and convex subset of $B(\ell^q(\mathbb{N}^l))$ and hence is weakly compact by the previous argument. So by Theorem 3.1 it suffices to show that $(W, d_q)/S_\mathbb{N}$ is totally bounded. To this end choose $\varepsilon > 0$. Let $k = k(p, q, \varepsilon/2)$ be the constant supplied by Lemma 5.3. Denote by $[n]$ the set of positive integers $\{1, \ldots, n\}$ for any $n \in \mathbb{N}$. Let F be a finite set of points in $B(\ell^q([k^l]^l))$ such that for each $y \in B(\ell^q([k^l]^l))$ there exists $x \in F$ with $\|x - y\|_q \le \varepsilon/2$. By Lemma 5.3, for each $w \in W$ there exists a set $K \subset \mathbb{N}^l$ of size at most k such that $\|w_K - w\|_q \le \varepsilon/2$. Then there exists $\sigma \in S_\mathbb{N}$ such that $\sigma(K) \subseteq [k^l]^l$, as the l-tuples in K can contain at most k^l distinct numbers. This implies that $d_q(S_\mathbb{N} F, W) \le \varepsilon$ and finishes the proof. $\qquad\square$

References

[1] N. Alon, E. Fischer, M. Krivelevich and M. Szegedy: Efficient testing of large graphs, *Combinatorica* **20** (2000), 451–476.

[2] N. Alon, T. Lee, A. Shraibman and S. Vempala: The approximate rank of a matrix and its algorithmic applications, in: *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pp. 675–684. ACM, 2013.

[3] S. Barman: Approximating Nash Equilibria and Dense Subgraphs via an Approximate Version of Carathéodory's Theorem, arXiv preprint arXiv:1406.2296 (2015).

[4] J. Bergh and J. Löfström: *Interpolation Spaces An Introduction*, Grundlheren der mathematischen Wissenschaften **223**, Springer-Verlag Berlin Heidelberg New York, 1976.

[5] C. Borgs, J.T. Chayes, H. Cohn and Y. Zhao: An L^p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions, arXiv preprint arXiv:1401.2906 (2014).

[6] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, B. Szegedy and K. Vesztergombi: Graph limits and parameter testing, in: *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pp. 261–270. ACM, 2006.

[7] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós and K. Vesztergombi: Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing, *Advances in Mathematics* **219** (2008), 1801–1851.

[8] J.B. Conway: *A Course in Functional Analysis–Second Edition*, Graduate Texts in Mathematics, 96, Springer Science & Business Media, New York, 1990.

[9] J.N. Cooper: A permutation regularity lemma, *The Electronic Journal of Combinatorics* **13** (2006), R22.

[10] W. J. Davis, T. Fiegel, W. B. Johnson and A. Pelczynski: Factoring weakly compact operators, *J. Funct. Anal.* **17** (1974), 311–327.

[11] W. Fenandez de la Vega, M. Karpinski, R. Kannan and S. Vempala: Tensor decomposition and approximation schemes for constraint satisfaction problems, in *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pp. 747-754, ACM, 2005.

[12] A. Frieze and R. Kannan: Quick approximation to matrices and applications, *Combinatorica* **19** (1999), 175-220.

[13] B. Green: A Szemerédi-type regularity lemma in abelian groups, with applications, *Geometric and Functional Analysis* **15** (2005), 340–376.

[14] A. Grothendieck, Critères de compacité dans les espaces fonctionelles généraux, *Amer. J. Math.* **74** (1952), 168–186.

[15] P.R. Halmos: *Measure theory*, Graduate Texts in Mathematics, 18, Springer, New York, 2014.

[16] S. Janson: Graphons, cut norm and distance, couplings and rearrangements, *NYJM Monographs* Vol 4, 2013.

[17] Y. Kohayakawa: Szemerédi's regularity lemma for sparse graphs, in: *Foundations of computational mathematics*, pp. 216–230, Springer Berlin Heidelberg, 1997.

[18] J. Komlós and M. Simonovits: Szemerédi's Regularity Lemma and its applications in graph theory, *Combinatorics, Paul Erdos is Eighty*, (D. Miklos et. al, eds.), Bolyai Society Mathematical Studies, 2, 295–352, 1996.

[19] L. Lovász and B. Szegedy: Szemerédi's Lemma for the analyst, *Geometric and Functional Analysis* **17** (2007), 252–270.

[20] L. Lovász and B. Szegedy: Limits of dense graph sequences, *Journal of Combinatorial Theory, Series B* **96** (2006), 933–957.

[21] L. Lovász: *Large Networks and Graph Limits*, American Mathematical Society, Providence, Rhode Island, 2012.

[22] T. Nishiura: Measure-Preserving Maps of $\mathbb{R}^n$, *Real Analysis Exchange* **24** (1999), 837–842.

[23] G. Regts: Approximation algorithms for matrix norms from regularity lemmas in Banach spaces, in preparation.

[24] G. Regts and A. Schrijver: Compact orbit spaces in Hilbert spaces and limits of edge-colouring models, to appear in *European Journal of Combinatorics*, arXiv preprint, arXiv.1210.2204 (2012).

[25] V. Rödl and J. Skokan: Regularity Lemma for k-uniform hypergraphs, *Random Structures and Algorithms* **25** (2004), 1-42.

[26] A. Scott: Szemerédi's regularity lemma for matrices and sparse graphs, *Combinatorics, Probability and Computing* **20** (2011), 455–466.

[27] E. Szemerédi: Regular partitions of graphs, in: Problèmes combinatoires et théorie des graphes, (Proceedings Colloque International C.N.R.S., Paris-Orsay, 1976) [Colloques Internationaux du C.N.R.S. No 260], Editions du C.N.R.S., Paris, 1978, pp. 399–401.

[28] T.C. Tao: Szemerédi's regularity lemma revisited, *Contributions Discrete Math.* **1** (2006), 8–28.